

5 What is Claimed is:

1. A method for categorizing voice samples of a person being tested for near term suicidal risk as a prelude to such testing, comprising the steps of:

10 A. setting an analysis window to a selected sample set length of 512, where the particular sample is identified as the Kth sample;

B. reading the Kth sample;

C. computing wavelet transforms of such Kth sample for scales in powers of 2 running from the 1st power to the 5th;

D. storing the signal energy value as computed for each scale;

15 E. checking to determine whether the Kth sample is the last of the sample set and if additional samples remain, repeating steps "b" through "d";

F. setting the median energy distribution at the scale for 2 to the 4th power as a threshold;

20 G. successively for each sample comparing the energy across the scales;

H. if the maximum energy is at the scale for 2 to the 1st power, identifying the segment as unvoiced and proceeding to the next succeeding sample;

25 I. if the segment maximum energy is at one of the scales of 2 to the 2nd power through 2 to the 5th power, identifying the segment as being either voiced or silence; and

J. if the segment energy at the 2 to the 4th power scale exceeds the threshold, classifying the segment as voiced; otherwise classifying it as silence.

30 2. A method for determining jitter variations in fundamental frequency of the voice of a person being evaluated for near-term suicidal risk, comprising the steps of:

5 A. setting an analysis window to a selected sample set length of 512 where the particular sample is identified as the Kth sample;

10 B. computing the wavelet transform for the sample set at scale 2 to the 4th power, with a scale factor defined by the quotient of the wavelet center frequency at level 0 and the desired center frequency;

15 C. selecting two consecutive segments of the vocal signal of such person which are voiced segments and generating separate pulse trains in which the heights of the pulses correspond to amplitude of positive and negative peaks of the wavelet transformed speech signal;

20 D. thresholding the segments of the vocal signal to discard peaks corresponding to possible unvoiced samples;

25 E. computing a fundamental period over the entirety of each of the two segments by:

30 i. finding the location of the first peak of the autocorrelation of the smoothed spectrum to the right of the zero lag component;

 ii. detecting a starting pulse exhibiting the property of being larger than both the pulse immediately preceding and immediately following such pulse and being greater than 50% of the global maximum of the pulse sequence;

 iii. locating following prominent pulses as detected in the neighborhood of expected locations determined by the peak of the autocorrelation sequence;

 iv. selecting, between two sequences of positive and negative peaks, the peak having the largest magnitude; and

 v. taking the difference between two consecutive prominent pulses as the duration for the glottal cycle; and

35 F. determining period-to-period fluctuation of fundamental frequency as the inverse of said glottal cycle for said two consecutive prominent pulses.

5 3. A method for testing voice samples of a person for near-term suicidal risk, comprising the steps of:

10 A. setting an analysis window to a selected sample set length of 512 consecutive segment voice signals from the person, where the particular sample is identified as the Kth sample;

15 B. reading the Kth sample;

15 C. computing wavelet transforms of such Kth sample for scales in powers of 2 running from the 1st power to the 5th;

20 D. storing the signal energy value as computed for each scale;

20 E. checking to determine whether the Kth sample is the last of the sample set and if additional samples remain, repeating steps "b" through "d";

25 F. setting the median energy distribution for the scale at 2 to the 4th power as a threshold;

25 G. successively for each sample comparing the energy across the scales;

30 H. if the maximum energy is at the scale for 2 to the 1st power, identifying the segment as unvoiced and proceeding to the next succeeding sample;

30 I. if the segment maximum energy is at one of the scales 2 to the 2nd power through 2 to the 5th power, identifying the segment as being either voiced or silence;

30 J. if the segment energy at the 2 to the 4th power scale exceeds the threshold, classifying the segment as voiced; otherwise classifying it as silence;

30 K. computing the wavelet transform for the sample set at scale 2 to the 4th power, with a scale factor defined by the quotient of the wavelet center frequency at level 0 and the desired center frequency;

30 L. selecting two consecutive segments of said vocal signal of such person which have been identified as voiced segments and generating two

5 separate pulse trains in which heights of pulses correspond to amplitude of positive and negative peaks of the wavelet transformed speech signal;

10 M. thresholding the segments of the vocal signal to discard peaks corresponding to possible unvoiced samples;

15 N. computing a fundamental frequency over the entirety of each of the two segments by

20 i. finding the location of the first peak of the autocorrelation of the smoothed spectrum to the right of the zero lag component;

25 ii. detecting a starting pulse exhibiting the property of being larger than the pulse immediately preceding and immediately following such pulse and being greater than 50% of the global maximum of the pulse sequence;

30 iii. locating following prominent pulses as detected in the neighborhood of expected locations determined by the peak of the autocorrelation sequence;

35 iv. selecting between two sequences of positive and negative pulse peaks the peak having the largest magnitude; and

40 v. taking the difference between two consecutive prominent pulses as the duration for the glottal cycle; and

45 O. determining period-to-period fluctuation of fundamental frequency as the inverse of said glottal cycle for said two consecutive prominent pulses.

4. A method for assessing near-term suicidal risk through voice analysis independently of verbal content of the voice, comprising:

30 A. eliciting a voice sample from a person to be evaluated for near-term suicidal risk and converting said sample into electronically processable signal form;

35 B. time-wise dividing said signal into segments according to whether the person was silent, speaking voiced words or making unintelligible unvoiced sounds;

5 C. if there are two consecutive voiced segments, measuring fundamental frequency for each of said two segments;

10 D. computing the difference in measured fundamental frequency for said two segments;

15 E. comparing the difference in measured fundamental frequency to fundamental frequency difference data for known near-term suicidal risk persons, known depressed persons not at near-term suicidal risk and non-depressed persons from a control group, to determine whether the person is at near-term suicidal risk or is merely depressed.

20 5. Apparatus for assessing near-term suicidal risk through voice analysis independently of verbal content of an elicited vocal utterance from a person to be evaluated for near-term suicidal risk, comprising:

25 A. means for converting said utterance into electronically processable signal form and time-wise dividing said signal into segments according to whether the person was silent, speaking voiced words or making unintelligible unvoiced sounds;

30 B. comparator means for determining whether there are two consecutive voiced segments and, if so, measuring fundamental frequency for each of said two segments;

35 C. means for computing difference in measured fundamental frequency for said two segments; and

40 D. means for comparing difference in measured fundamental frequency to fundamental frequency difference data for known high near-term suicidal risk persons, known depressed persons not at near-term suicidal risk and non-depressed persons from a control group, to determine whether the person is at near-term suicidal risk or is merely depressed, and providing a visual and/or audible alarm signal upon finding near-term suicidal risk.

5 6. Apparatus for evaluating whether a person is at high near term risk of
suicide by determining jitter variations in fundamental frequency of the voice of
such a person being evaluated for near-term suicidal risk, comprising:

10 A. counter means for setting an analysis window to a selected sample
set length of 512 where the particular sample is identified as the Kth
sample and computing the wavelet transform for the sample set at scale 2
to the 4th power, with a scale factor defined by the quotient of the wavelet
center frequency at level 0 and the desired center frequency;

15 B. means for selecting two consecutive segments of the vocal signal
of such person which are voiced segments and generating two separate
pulse trains in which the heights of the pulses correspond to the amplitude
of positive and negative peaks of the wavelet transformed speech signal,
thresholding the segments of the vocal signal to discard peaks
corresponding to possible unvoiced samples and computing a fundamental
period over the entirety of each of the two segments by:

20 i. finding the location of the first peak of the autocorrelation
of the smoothed spectrum to the right of the zero lag component;
ii. detecting a starting pulse exhibiting the property of being
larger than the pulse immediately preceding and immediately
following such pulse and being greater than 50% of the global
maximum of the pulse sequence;

25 iii. locating following prominent pulses as detected in the
neighborhood of expected locations determined by the peak of the
autocorrelation sequence;
iv. selecting, between two sequences of positive and negative
peaks, the peak having the largest absolute magnitude; and
v. taking the difference between two consecutive prominent
pulses as the duration for the glottal cycle; and

30 C. means for determining period-to-period fluctuation of fundamental
frequency as the inverse of said glottal cycle for said two consecutive
prominent pulses, comparing said fundamental frequency fluctuation for

35

5 such person to that of persons known to be not at near term risk for suicide and providing a clinician with a visual and/or audible signal indicating the results of such comparison.

10 7. Apparatus for segregating into categories voice samples of persons being tested for near term suicidal risk as a prelude to such testing, comprising:

15 A. counting means for setting an analysis window to a selected sample set length of 512, where the particular sample is identified as the Kth sample, and reading the Kth sample;

B. means for computing wavelet transforms of such Kth sample for scales in powers of 2 running from the 1st power to the 5th and storing signal energy values as computed for each scale;

C. means for checking to determine whether the Kth sample is the last of the sample set and, if additional samples remain, triggering said computing means to repeat;

D. storage means for maintaining a median energy distribution at scale 2 to the 4th power as a threshold;

E. comparator means for comparing energy across the scales for each sample successively and if maximum energy is at the scale for 2 to the 1st power, identifying the segment as unvoiced and proceeding to the next succeeding sample but if the segment maximum energy is at one of the scales 2 to the 2nd power through 2 to the 5th power, identifying the segment as being either voiced or silence and further if segment energy at the 2 to the 4th power scale exceeds the threshold, identifying the segment as voiced; otherwise identifying it as silence.

20 30 8. Apparatus for testing voice samples of persons for near-term suicidal risk and providing an alarm signal upon a subject being found to represent a near-term suicidal risk, comprising:

35 A. counter means for setting an analysis window to a selected sample set length of 512 consecutive segment voice signals from the person,

5 where the particular sample is identified as the Kth sample, and reading
the Kth sample of said person's vocal signal;

10 B. mathematical processor means for computing wavelet transforms
of such Kth sample for scales in powers of 2 running from the 1st power
to the 5th and storing the signal energy value as computed for each scale
while checking to determine whether the Kth sample is the last of the
sample set and if additional samples remain, repeating the wavelet
transformation;

15 C. storage means for setting and storing median energy distribution at
the scale for 2 to the 4th power as a threshold;

20 D. comparator means for successively comparing, for each sample,
energy across the scales and testing to determine if the maximum energy is
at the scale for 2 to the 1st power and identifying the segment as unvoiced
and proceeding to the next succeeding sample but if the segment
maximum energy is at one of the scales 2 to the 2nd power through 2 to the
5th power, identifying the segment as being either voiced or silence and
further comparing segment energy at the 2 to the 4th power scale for
exceeding the threshold and identifying the segment as voiced; otherwise
identifying it as silence;

25 E. means for computing the wavelet transform for the sample set at
scale 2 to the 4th power, with a scale factor defined by the quotient of the
wavelet center frequency at level 0 and the desired center frequency and
selecting two consecutive segments of said vocal signal of such person
identified as voiced segments and generating separate pulse trains in
which heights of pulses correspond to amplitude of positive and negative
peaks of the wavelet transformed speech signal;

30 F. peak detector means for thresholding the segments of the vocal
signal and discarding peaks corresponding to possible unvoiced samples;

G. processor means for computing fundamental period over the
entirety of each of the two segments by :

5

i. finding the location of the first peak of the autocorrelation of the smoothed spectrum to the right of the zero lag component;

10

ii. detecting a starting pulse exhibiting the property of being larger than the pulse immediately preceding and immediately following such pulse and being greater than 50% of the global maximum of the pulse sequence;

15

iii. locating following prominent pulses as detected in the neighborhood of expected locations determined by the peak of the autocorrelation sequence;

20

iv. selecting between two sequences of positive and negative peaks the peak having the largest peak magnitude; and

H. means for taking the difference between two consecutive prominent pulses as the duration for the glottal cycle and determining period-to-period fluctuation of fundamental frequency as the inverse of said glottal cycle for said two consecutive prominent pulses.

RECORDED BY: JAMES R. HARRIS
APPROVED BY: JAMES R. HARRIS
DATE: 10/05/01
TIME: 11:37 AM